

基于渲染图像角度结构特征的三维模型检索方法

刘 志 潘晓彬

(浙江工业大学计算机科学与技术学院 杭州 310023)

摘 要 为了充分利用三维模型的颜色、形状、纹理等特征,提出以三维模型渲染图像为数据集,利用渲染图像角度结构特征实现三维模型检索。首先,该方法以三维模型渲染图像为测试集,利用已有类别标记的自然图像作为训练集,通过骨架形状上下文特征对渲染图像进行分类,提取角度结构特征,建立特征库;然后,对输入的自然图像提取角度结构特征,与特征库中的角度结构特征进行相似度匹配计算,实现三维模型检索。实验结果表明,充分利用渲染图像的颜色、形状和空间信息是实现三维模型检索的有效方法。

关键词 三维模型检索,渲染图像,骨架形状上下文,角度结构特征

中图法分类号 TP391 文献标识码 A

3D Model Retrieval Method Based on Angle Structure Feature of Render Image

LIU Zhi PAN Xiao-bin

(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract In order to make full use of the color, shape, texture and other features in the 3D model, a 3D model retrieval method was proposed based on angle structure features of render images. Firstly, the 3D model render images are taken as a test dataset and the marked natural images are taken as a training set. The render images are classified based on their skeleton-associated shape context and the angle structure features are extracted to establish the feature library. Then, the angle structure features of the input natural images are extracted. The distance measurement method is used to calculate the similarity between the angle structure feature of input natural image and those features in the feature library. The experimental results show that the full utilization of the color, shape and color space information of the render image is an effective way to achieve 3D model retrieval.

Keywords 3D model retrieval, Render image, Skeleton-associated shape context, Angle structure feature

1 引言

三维模型作为继声音、图像和视频之后的第四代多媒体数据类型,已经被广泛应用于工业产品设计、虚拟现实、3D 游戏和影视动画等众多领域。随着三维模型数量的不断增多,如何更快、更好地从大量的已有三维模型中检索出与用户需求最相似的三维模型变得至关重要。

目前,基于内容的三维模型检索的主要方法有基于草图、二维投影视图和示例模型的三维模型检索。在基于草图和二维投影视图的相关研究中,Yoon 等^[1]提出对输入草图和投影得到的暗示轮廓使用扩散张量场进行特征提取,采用方向梯度直方图实现相似度计算;Lei 等^[2]提出集成草图和投影视图的骨架图和轮廓特征来实现检索。这两种方法都侧重于利用草图和视图的形状信息,对纹理、颜色等信息的应用不足。

广泛应用于图形图像检索方法中的颜色、纹理、形状等视觉信息也被应用于三维模型检索中^[3-6]。Pasqualotto 等^[7]提出使用网格图像描述符获取三维模型的形状和颜色两种视觉信息来实现检索,在使用三维模型形状信息的同时应用了颜

色信息。Biasotti 等^[8]对 6 种纹理三维形状检索和分类方法进行了分析和比较,通过实验验证了组合使用几何形状、纹理和颜色 3 种视觉信息实现三维模型检索的可行性。

文献^[9]提出使用微观结构描述符对局部图像块提取相似边缘取向的底层微观结构,并以此为桥梁集图像颜色、纹理、形状和颜色分布信息为一体以描述图像。文献^[10]提出使用显著结构直方图描述图像。文献^[11]提出角度结构特征概念,即对图像进行颜色量化后,通过统计所有局部块中各邻近像素点间的角度及大小关系数量来实现相似度计算。在图像检索过程中,角度结构特征具有较强的识别能力。

由于三维模型渲染图像充分反映了三维模型的颜色、形状、空间等特征信息,本文提出了一种基于渲染图像角度结构特征的三维模型检索方法。该方法以分类后的三维模型渲染图像为数据集,通过数据集渲染图像的角度结构特征库,与输入自然图像进行相似性匹配,实现三维模型检索。

2 算法框架

本文方法分为两个部分:1)获取三维模型渲染图像作为

本文受浙江省自然科学基金(LY16F020033)资助。

刘 志(1969—),女,博士,教授,CCF 会员,主要研究方向为三维模型检索、图像处理等,E-mail: lzhi@zjut.edu.cn;潘晓彬(1992—),男,硕士生,主要研究方向为三维模型检索。

测试集,使用已有类别标记的自然图像作为训练集,基于骨架形状上下文特征训练一个多类别线性 SVM 分类器从而实现渲染图像分类,形成数据集,并提取角度结构特征,建立特征库;2)对输入的自然图像提取角度结构特征,基于距离度量方法实现三维模型检索。算法框架如图 1 所示。

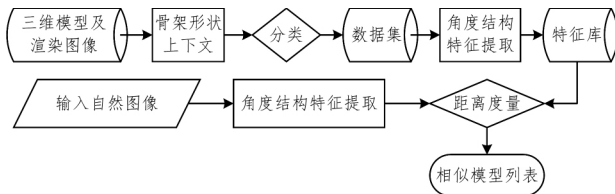


图 1 算法框架

3 渲染图像分类

本节主要通过骨架形状上下文特征^[12]对渲染图像实现分类。首先对渲染图像进行二值化(黑色背景、白色前景),利用 Shen 等^[13]的方法提取图像骨架,然后为每个轮廓点寻找最近骨架点,并将两点间的距离值赋给轮廓点,完成轮廓与骨架信息的结合。然后分段计算骨架形状上下文,并编码分段的骨架上下文形成特征向量。最后获取已有类别标记的自然图像作为训练集,基于骨架形状上下文特征训练一个多类别线性 SVM^[14-15]作为分类器,用于渲染图像分类,以达到三维模型分类的效果。

3.1 主干骨架图提取

在提取渲染图像的骨架图前,首先需要对其进行二值化预处理。而渲染图像(见图 2(a))进行初步二值化后得到的二值图像(见图 2(b))还是会出现部分阴影无法去除的情况。为了得到较优的骨架图,需要对图 2(b)所示的二值图像进行二次腐蚀和膨胀,消除多余阴影,从而得到较优的二值图像(见图 2(c))。



图 2 二值图像优化

Shen 等^[13]提出的提取骨架的方法,是在常用的骨架算法获得骨架图(见图 3(a))的基础上,使用式(1)计算骨架的每段末端分支的权重,多次迭代删除权重最小的末端分支,达到修剪骨架的目的,从而形成主干的骨架图(见图 3(b)),使之能够更好地表达图像的形状特性。

$$w_j^i = \alpha \frac{\Lambda(D - R(S^i - E_j^i))}{\Lambda(D)} + \log(\Gamma(S^i - E_j^i) + 1) \quad (1)$$

$$\alpha = \beta \log\left(\frac{\Gamma(M)}{\bar{\omega}(M)}\right), S^0 = M, S^{i+1} = S^i - E_i \quad (2)$$

其中, D 是形状轮廓点集; S 是骨架点,它的上标表示迭代次数; E 是末端分支节点,它的上标与 S 的上标的意义相同,下标表示末端分支序号; α 是归一化因数; β 定义为常量 9; $\bar{\omega}(M)$ 表示末端路径 M 的平均长度; R 表示重构骨架点(即将一个末端分支删除后的骨架点); Γ 表示归一化的曲线长度; Λ 表示面积以像素为单位。

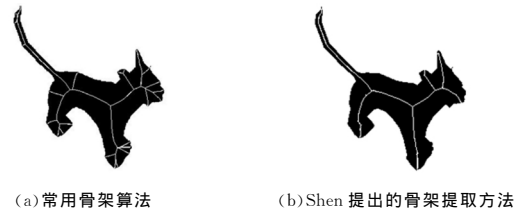


图 3 骨架图优化

3.2 骨架形状上下文特征提取

骨架形状上下文是在图像轮廓点的形状上下文的基础上,增加骨架信息而得到(轮廓点不仅有位置信息,还有该轮廓点与最近的骨架点间的距离信息)。最终形成的直方图从形状上下文的 60 维(12 个数字用来表示角度大小,5 个数字用来表示半径)变成骨架形状上下文的 300 维(12 个数字用来表示角度大小,5 个数字用来表示半径大小,5 个数字用来表示骨架值)。然后将包含骨架信息的轮廓点分段,求出每段中 5 个轮廓点间的骨架形状上下文,形成 1500 维的骨架形状上下文特征包,最后对每段的特征包进行编码,得到骨架形状上下文特征向量集,如图 4 所示。



图 4 骨架形状上下文

3.3 渲染图像分类

对获取的自然图像进行类别标记后,将其作为训练集,通过 3.2 节的方法获得训练集的骨架形状上下文特征向量集,然后训练一个多类别线性 SVM 作为分类器,如式(3)所示:

$$\min_{w_1, \dots, w_L} \sum_{j=1}^M \|w_j\|^2 + \gamma \sum_i \max(0, 1 + w_{l_i}^T g_i - w_{y_i}^T g_i) \quad (3)$$

$$l_i = \arg \max_{l \in \{1, 2, \dots, L\}, l \neq y_i} w_l^T g_i \quad (4)$$

式(3)中的第一项是正则项,用来最大化划分超平面的间隔;第二项是 hinge 损失函数,用来衡量预测器对输入数据进行类别预测时的质量好坏; γ 是用于平衡正则项(左部)和 hinge 损失函数(右部)的正则化常数,定义为常量 10; M 表示所有类别的训练图像; g_i 是训练图像的特征向量集; y_i 是训练图像类别标签。

将三维模型的渲染图像作为测试集,通过 3.2 节的方法获得测试集的骨架形状上下文特征向量集,然后根据分类器

和预测函数(见式(5))获得测试图像的分类标签 \hat{y} 。根据测试集的分类标签,将同一类别的渲染图像和三维模型移动到同一文件夹下,间接完成三维模型的分类,建立数据集。

$$\hat{y} = \arg \max_{i \in \{1, 2, \dots, L\}} w_i^T g \quad (5)$$

其中, g 是测试图像的特征向量。

4 三维模型检索

首先对数据集中所有的渲染图像进行颜色量化;然后提取渲染图像的角度结构特征,建立特征库;最后对输入图像提取角度结构特征,采用距离度量方法从特征库中检索相似模型,实现三维模型检索。

4.1 渲染图像颜色量化

为了降低计算复杂度和更好地描述图像的颜色和形状信息,HSV 颜色空间被量化为 72 个空间^[18],如图 5 所示。

$$H = \begin{cases} 0, & H \in [0, 24] \cup [345, 360] \\ 1, & H \in [25, 49] \\ 2, & H \in [50, 79] \\ 3, & H \in [80, 159] \\ 4, & H \in [160, 194] \\ 5, & H \in [195, 264] \\ 6, & H \in [265, 284] \\ 7, & H \in [285, 344] \end{cases}$$

$$S = \begin{cases} 0, & S \in [0, 0.15] \\ 1, & S \in (0.15, 0.8] \\ 2, & S \in (0.8, 1] \end{cases}, V = \begin{cases} 0, & V \in [0, 0.15] \\ 1, & V \in (0.15, 0.8] \\ 2, & V \in (0.8, 1] \end{cases}$$

图 5 HSV 颜色空间量化

图 5 中, H 通道被划分为 8 个区域, S 通道被划分为 3 个区域, V 通道被划分为 3 个区域。

根据 S 和 V 两个通道划分情况的不同,得到的量化函数不同,如式(6)所示:

$$P = Q_S Q_V H + Q_V S + V \quad (6)$$

其中, Q_S 是颜色空间成分 S 的数字化的数量, Q_V 是颜色空间成分 V 的数字化的数量。本文中 S 和 V 被量化为 3 层,因此 $Q_S = 3, Q_V = 3$ 。最终得到的量化函数如式(7)所示:

$$P = 9H + 3S + V \quad (7)$$

然后使用量化函数,将图像每个像素的 H, S, V 3 个通道合并为一个值,最终图像成为了一个二维矩阵,矩阵中的每个值的位置都与图像的像素点位置逐一对应,矩阵每个值的大小为图像像素点通过量化函数得到的值。

假设 I 是图像进行颜色量化前的结果, i 被定义为颜色量化后的值, L_i 是 I 颜色量化后的值为 i 的点集,如式(8)所示:

$$L_i = \{(x, y) | (x, y) \in I, I(x, y) = i, 0 \leq i \leq 71\} \quad (8)$$

4.2 渲染图像角度结构特征的提取

图像角度结构特征的提取主要分为以下 3 个步骤:

1) 对于每一个量化后的颜色 α , 从坐标 $(0, 0)$ 开始, 从左到右、从上到下移动 2×2 的区域去遍历量化后的颜色矩阵。在遍历的过程中, 颜色矩阵被分割为很多个 2×2 的局部块。对于每一个局部块来说, 如果 V_1 和当前量化后的颜色 α 相等, 则保留这个局部块, 否则移除这个局部块。

2) 对于保留的局部块, 我们获取同一个块中 V_1 与其他像素 V_2, V_3, V_4 的角度关系(见图 6), 以及同一个块中 V_1 和

其他像素 V_2, V_3, V_4 的大小关系(见图 7)。

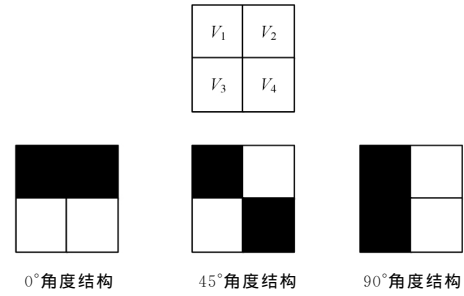


图 6 角度关系

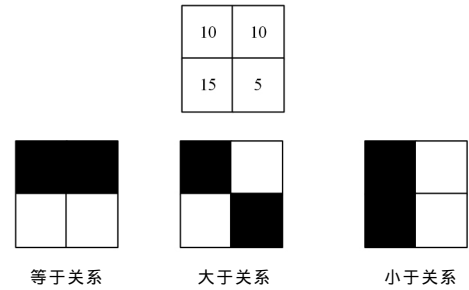


图 7 大小关系

统计所有角度下的大小关系后, 得到一个描述颜色空间信息的 9 维向量。其中, N_0^E, N_0^L, N_0^G 是在 0° 的角度结构下的等于、小于和大于关系的数量。同样地, $N_{45^\circ}^E, N_{45^\circ}^L, N_{45^\circ}^G$ 和 $N_{90^\circ}^E, N_{90^\circ}^L, N_{90^\circ}^G$ 分别是在 $45^\circ, 90^\circ$ 角度结构下的大小关系的数量。例如, 以图 8 所示图像为原始图像, 可以得到如下的一个 9 维的向量(假设当前量化后的颜色值为 10)。

$$(N_0^E, N_0^L, N_0^G, N_{45^\circ}^E, N_{45^\circ}^L, N_{45^\circ}^G, N_{90^\circ}^E, N_{90^\circ}^L, N_{90^\circ}^G) = (1, 0, 2, 1, 1, 1, 2, 1, 0)$$

10	5	10	10
15	10	10	5
		10	5
		10	15

图 8 原始图像

3) 为了降低向量维数, 定义 $T = (t_1, t_2, t_3)$ 表示角度结构, 其中 t_1 是大于关系, t_2 是小于关系, t_3 是等于关系。 $T^E = (0, 0, 1), T^L = (0, 1, 0), T^G = (1, 0, 0)$ 分别表示等于关系、小于关系和大于关系。 T 如式(9)所示:

$$T = \sum_{i=1}^3 t_i 2^{(i-1)} \quad (9)$$

定义 $L_{\alpha_0^\circ}, L_{\alpha_{45^\circ}}, L_{\alpha_{90^\circ}}$ 为当前量化后的颜色值为 α 的 3 个角度结构的值, 以 0° 角度结构为例, 通过式(10)计算得到相应的值。

$$L_{\alpha_0^\circ} = T^E N_0^E + T^L N_0^L + T^G N_0^G \quad (10)$$

最终, 可以得到 216 维的图像角度结构特征向量 H :

$$H = [L_{0^\circ}, L_{0_{45^\circ}}, L_{0_{90^\circ}}, \dots, L_{\alpha_0^\circ}, L_{\alpha_{45^\circ}}, L_{\alpha_{90^\circ}}, \dots, L_{71_{45^\circ}}, L_{71_{90^\circ}}]$$

最后通过以上方法, 提取数据集中渲染图像的角度结构特征, 建立特征库。

4.3 相似度计算

对输入的自然图像提取相同的角度结构特征, 采用距离度量方法从特征库中检索相似模型, 并以相似度大小排序。

本节分别使用曼哈顿距离、欧几里德距离和文献[11]提出的距离度量方法来进行相似度计算,得到在检索结果为 50 个的情况下的平均查准率,如表 1 所列。通过实验对比,本文最终采用文献[11]提出的距离度量方法。

表 1 不同距离下的 AP

距离度量方法	AP
本文距离度量方法	0.371
曼哈顿距离	0.324
欧几里德距离	0.323

假设输入图像的特征向量 $Q=[Q_1, Q_2, \dots, Q_L]$, 数据集中的图像的特征向量 $T=[T_1, T_2, \dots, T_L]$, 两者的距离如式(11)所示:

$$D(T, Q) = \sum_{i=1}^L \frac{|T_i - Q_i|}{1 + T_i + Q_i} \quad (11)$$

5 实验分析

本节对提出的方法进行检索实验及性能比较分析。本文方法在 Windows 平台上实现。硬件配置为 CPU Intel Core™ i5-2450M, 内存为 8 GB。实验采用的三维模型数据集是从 3dwarehouse, sketchup, com, artist-3d, com 和 archive3d, net 上根据文本后缀与关键词获取和挑选的包含自然物体属性并具有良好网格密度的三维网格模型得到的, 同时获取了相应的渲染图像。

5.1 实验设置

本文参考文献[19]提出的使用 17 个不同类别的模型进行实验的思路, 将实验的模型类别限定为四肢动物、车、杯子、盆栽、人物、吉他等 11 个类别。本文使用互联网上的四肢动物、车、杯子、盆栽、人物、吉他等 11 个类别的各 40 张自然图像作为训练集, 从 3 个网站上获取的四肢动物、车、杯子、盆栽、人物、吉他等 11 个类别的各 80 个三维模型及其渲染图像作为测试集, 通过训练集的骨架形状上下文特征训练出多类别线性 SVM 作为分类器, 对测试集渲染图像进行分类; 从分类后的每个类别中分别选取 40 张渲染图像作为数据集, 然后使用网上获得的四肢动物、车、杯子、盆栽、人物、吉他等 11 个类别的各 20 张自然图像作为输入图像, 使用 216 维的角度结构特征进行图像检索, 从而实现三维模型检索。数据集中所有渲染图像为 128×128 或者 230×130 的 jpg 格式, 所有三维模型类型包括 obj, off 和 skp 3 种格式。

5.2 实验验证分析

在本实验中, 首先对渲染图像进行颜色量化, 以此来量化描述图像的颜色和形状信息, 然后利用角度结构将颜色、形状和颜色空间信息集成为一体, 建立 216 维的角度结构特征向量, 最后通过距离度量方法比较输入图像和数据集渲染图像间的向量差异, 并对差异大小进行排序, 实现三维模型检索。本文从四肢动物、车和吉他 3 个类别中分别选取了一张输入图像, 然后检索出前 12 个结果, 如图 9 所示。

从图 9 的检索结果可知, 对于不同类别、形态的三维模型, 以包含较多视觉信息的自然图像为输入源, 本文方法具有较好的检索表现。

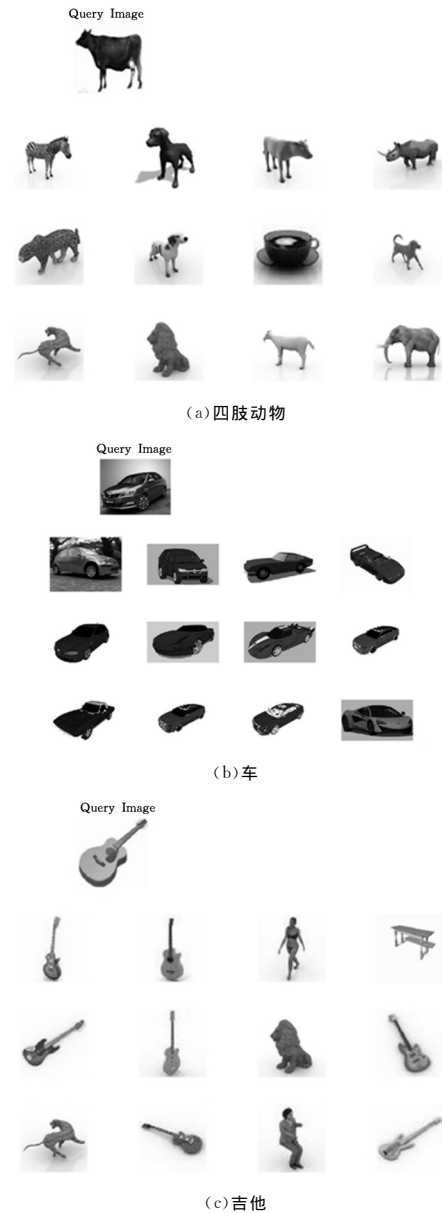


图 9 采用本文方法的三维模型检索示例

5.3 对比实验

在实验数据相同的 9 个类别下, 使用第一等级匹配 (First-Tier, FT) 评价指标与 Gabor+BOW^[19] 进行横向对比, 结果如图 10 所示。

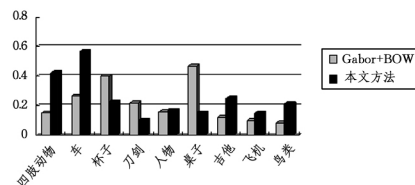


图 10 部分类别的 FT 实验对比

从对比情况可以看出, 在四肢动物、车、吉他和鸟类等较为复杂的类别上, 本文采用的检索方法具有较好的检索性能。但是在杯子、刀剑和桌子等简单类别上, 由于对应的自然图像及渲染图像在邻近像素点间的角度结构关系变化比较简单, 提取到的视觉信息较少, 形成的角度结构特征的区分力下降, 对检索结果产生了较大影响, 这也在一定程度上说明本文方法在部分类别中存在优化空间。

为了使检索评价更加全面,以第二等级匹配(Second-Tier, ST)、E 指数(E-Measures, EM)、最近邻域(Nearest Neighbor, NN)、标准化多值相关度(Normalized Discounted Cumulative Gain, nDCG)和平均查准率(Average Precision, AP)^[20] 5 个评价指标为依据,量化对比实验数据如表 2 所列。

表 2 检索性能对比表

评价指标	本文方法	Gabor+BOW ^[19]
ST	0.374	0.370
EM	0.247	0.239
nDCG	0.267	0.265
NN	0.436	0.355
AP	0.371	0.357

结束语 本文提出以三维模型渲染图像为数据集,利用渲染图像角度结构特征实现三维模型检索。该方法通过建立骨架形状上下文特征分类器,实现了对渲染图像的分类学习,利用距离相似度计算,实现了对来自三维模型的自然图像的相似性匹配,进而实现了三维模型检索。将广泛应用于二维图像的检索方法应用于三维模型检索,具有良好的检索效果,尤其是对于渲染图像邻近像素点间的角度结构关系变化复杂的三维模型,具有较好的检索效果。但受限于渲染图像的数量和质量,其普适性也受到了一定的影响。

参 考 文 献

- [1] YOON S M, SCHERER M, SCHRECK T, et al. Sketch-based 3D model retrieval using diffusion tensor fields of suggestive contours[C]//Proceedings of the 18th ACM International Conference on Multimedia. ACM, 2010:193-200.
- [2] LEI H, LI Y, CHEN H, et al. A novel sketch-based 3D model retrieval method by integrating skeleton graph and contour feature[J]. Journal of Advanced Mechanical Design, Systems, and Manufacturing, 2015, 9(4).
- [3] ZHANG J, BAI C, NEZAN J F, et al. Joint motion model for local stereo video-matching method[J]. Optical Engineering, 2015, 54(12):123108.
- [4] ZHANG J, SHANG J, ZHANG G. Verification for different contrail parameterizations based on integrated satellite observation and ECMWF reanalysis data[J]. Advances in Meteorology, 2017, 2017(1):1-11.
- [5] 李兰, 刘洋. 基于内容的小波变换图像检索方法[J]. 计算机科学, 2015, 42(2):306-310.
- [6] 翟奥博, 温显斌, 张鑫. 基于改进双树复小波和灰度-梯度共生矩阵的纹理图像检索算法[J]. 计算机科学, 2017(6):274-277.
- [7] PASQUALOTTO G, ZANUTTIGH P, CORTELAZZO G M. Combining color and shape descriptors for 3D model retrieval[J]. Signal Processing: Image Communication, 2013, 28(6):608-623.
- [8] BIASOTTI S, CERRI A, AONO M, et al. Retrieval and classification methods for textured 3D models: a comparative study[J]. The Visual Computer, 2016, 32(2):217-241.
- [9] LIU G H, LI Z Y, ZHANG L, et al. Image retrieval based on micro-structure descriptor[J]. Pattern Recognition, 2011, 44(9):2123-2133.
- [10] LIU G H, YANG J Y, LI Z Y. Content-based image retrieval using computational visual attention model[J]. Pattern Recognition, 2015, 48(8):2554-2566.
- [11] ZHAO M, ZHANG H, MENG L. An angle structure descriptor for image retrieval[J]. China Communications, 2016, 13(8):222-230.
- [12] SHEN W, JIANG Y, GAO W, et al. Shape recognition by bag of skeleton-associated contour parts[J]. Pattern Recognition Letters, 2016, 83:321-329.
- [13] SHEN W, BAI X, YANG X W, et al. Skeleton pruning as trade-off between skeleton simplicity and reconstruction error[J]. Science China Information Sciences, 2013, 56(4):1-14.
- [14] CRAMMER K, SINGER Y. On the algorithmic implementation of multiclass kernel-based vector machines[J]. Journal of Machine Learning Research, 2001, 2(12):265-292.
- [15] MOORE R, DENERO J. L1 and L2 regularization for multiclass hinge loss models[C]//MLSLP. 2011:1-5.
- [16] BAI X, LIU W, TU Z. Integrating contour and skeleton for shape classification[C]//2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops). IEEE, 2009:360-367.
- [17] SHEN W, WANG X, YAO C, et al. Shape recognition by combining contour and skeleton into a mid-level representation[C]//Chinese Conference on Pattern Recognition. Springer Berlin Heidelberg, 2014:391-400.
- [18] WANG X, WANG Z. A novel method for image retrieval based on structure elements' descriptor[J]. Journal of Visual Communication and Image Representation, 2013, 24(1):63-74.
- [19] 刘志, 尹世超, 潘翔, 等. 基于特征线条的三维模型检索方法[J]. 计算机辅助设计与图形学学报, 2016, 28(9):1512-1520.
- [20] SHILANE P, MIN P, KAZHDAN M, et al. The princeton shape benchmark[C]//Proceedings of Shape Modeling Applications, 2004. IEEE, 2004:167-178.
- [6] MARTIN D R, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics [OL]. <https://wenku.baidu.com/view/c1ef906cb84ae45c3b358cfe.html>.
- [7] YANG A Y, WRIGHT J, MA Y, et al. Unsupervised segmentation of natural images via lossy data compression[J]. Computer Vision & Image Understanding, 2008, 110(2):212-225.
- [8] MA Y, DERKSEN H, HONG W, et al. Wright. Segmentation of multivariate mixed data via lossy coding and compression [J]. IEEE Transactions Pattern Analysis and Machine Intelligence, 2007, 29(9):1546-1562.
- [9] 李光, 王朝英, 侯志强. 基于 K 均值聚类与区域合并的彩色图像分割算法[J]. 计算机应用, 2010, 30(2):354-358.

(上接第 250 页)